SINTEF A22523- Unrestricted

# Report

## A survey of approaches for prioritizing trains in congested railroad networks

**Authors:**
Carlo Mannino
Adrian Werner
Arnt-Gunnar Lium


**SINTEF Technology and Society**
Applied Economics
2012-03-12

SINTEF Teknologi og samfunn
SINTEF Technology and Society

Address:
Postboks 4760 Sluppen
NO-7465 Trondheim
NORWAY

Telephone:+47 73593000
Telefax:+47 73590260

ts@sintef.no
www.sintef.no
Enterprise /VAT No:
NO 948007029 MVA

KEYWORDS:
Rail, prioritization of trains, infrastructure usage

# Report

# A survey of approaches for prioritizing trains in congested railroad networks

**VERSION**
1.0

**DATE**
2012-03-12

**AUTHOR(S)**
Carlo Mannino
Adrian Werner
Arnt-Gunnar Lium

**CLIENT(S)**
Regionalt Forskningsfond Hovedstaden

**CLIENT'S REF.**
Kjell Øygarden

**PROJECT NO.**
RFFHSTAD: 217659/SINTEF: 60A152

**NUMBER OF PAGES:**
15

**ABSTRACT**

On potentially highly congested railroad networks, demand for passenger and freight transport is affected not only by the available capacity but also by the consequences of the achieved punctuality and regularity: Poor on-time performance increases travel costs and, hence, makes rail less competitive toward other modes. It is, therefore, paramount to ensure a good punctuality record of the network, also in the case of disruptions and delays of single trains.

On the tactical level, well-planned timetables help to achieve both a high capacity utilization and good punctuality. On the operational level, advanced approaches for rescheduling or prioritizing trains can support dispatchers controlling the traffic flow.

In this report, we discuss issues to be considered when prioritizing trains and describe the most common approaches. Furthermore, we give a brief overview of some of the train dispatching systems currently in use and examples of where they are being used.

We also present some of the latest academic developments in more advanced decision support for train prioritization on the operational level. This also includes results of experimental studies for coordinating decisions over a larger part of the network. Finally, we discuss how to connect the dispatching of trains with incentive systems.

**PREPARED BY**
Arnt-Gunnar Lium

SIGNATURE
Arnt-Gunnar Lium

**CHECKED BY**
Nils Olsson
Nils Olsson

SIGNATURE

**APPROVED BY**
Frode Rømo

SIGNATURE
Frode Rømo

**REPORT NO.**
SINTEF A22523

**ISBN**
978-82-14-05-391-3

**CLASSIFICATION**
Unrestricted

**CLASSIFICATION THIS PAGE**
Unrestricted

# Document history

| VERSION | DATE | VERSION DESCRIPTION |
|---|---|---|
| 1.0 | 2012-03-12 | Version number 1.0 |

# Table of contents


**1 Introduction – Prioritizing trains in congested networks ..... 4**

**2 State of the art in prioritization methods ..... 6**

2.1 Prioritization methods applied in a real life setting ..... 6

Train dispatching systems currently in use ..... 7

2.2 State of the art in research ..... 9

2.3 Bridging the gap between theory and practice ..... 10

**3 Socio-economic issues and incentive systems ..... 10**

Combining the dispatching of trains and quality improving mechanisms ..... 11

Comparability to other modes of transportation ..... 12

**4 Conclusions ..... 13**

**References ..... 13**

# 1 Introduction – Prioritizing trains in congested networks

While passenger and freight transportation demand closely follows economic and population growth, the railways' share of the future traffic is expected to decline from 11 to 8 % for freight and from 6 to 5 % for passenger transport between 2000 and 2020 (Corman 2010). There are multiple reasons for this decrease; however, problems related to delays and frequency of service seem to play an important role. More passengers and freight migrating from rail to road would also require additional investments in road infrastructure in the future.

A report submitted to the European Commission (NEA 2003) concludes that the growth of the number of passengers using rail is, among other things, dependent on: 1) *removal of capacity restrictions of all types (infrastructure, stations, signalling and rolling stock)*, 2) *improved services (in terms of frequency or speed),* and *3) service performance and punctuality*. Hence, in order to stop the trend of railway's decreasing market share, more capacity and better services have to be offered. The Norwegian and the European railroad infrastructure is, in general, highly congested which creates problems for accommodating more passengers and freight. This can be addressed by investing in additional infrastructure, using the existing infrastructure (including rolling stock) better than today, or combining the two approaches. Building new and upgrading existing infrastructure is expected to have a positive impact on all the before-mentioned factors stimulating growth of rail passenger numbers. However, the expected cost of this will be high. According to Jernbaneverket (2012)[1], the cost of building or upgrading one kilometre of the Norwegian railroad infrastructure is expected to be in the range of about 200-400 million NOK[2] at the cost level of 2011. These estimates exclude risk premiums and costs for the principal.

The alternative to building additional infrastructure is to stimulate a more efficient utilization of existing and future infrastructure. One way of doing this is to operate the trains with a perfect or almost perfect on-time performance; this would greatly reduce or eliminate the need to set aside capacity (slack) in the network to deal with delays, unforeseen events, and other operational restrictions. Slack is usually inserted by assuming that trains will be driving slower than at optimal speed and/or by leaving some of the slots unutilized. If high on-time performance can be achieved then there will hardly be any difference between the planned timetable and the actually observed train operations.

Another way to accomplish higher utilization of the infrastructure is to support the dispatchers at the infrastructure manager Jernbaneverket. Enabling the dispatchers to make better operational decisions on the prioritization of trains through multiple bottlenecks will facilitate a higher utilization of existing assets such as tracks, stations, and rolling stock. This, in turn, allows the design of tighter timetables, increasing the utilization of the railway resources as well as improving on-time performance. Improving utilization of the infrastructure will not only stimulate railway services usage, it will also enhance the positive effects of investments into new infrastructure or upgrades of existing infrastructure.

Controlling and supervising the movements of trains is done by dispatchers at the infrastructure manager. When trains are running on schedule or in areas with few infrastructure capacity problems, they typically can follow a pre-programmed path through the network. Hence, in such situations, the dispatchers' task is just to execute the tactical plan (timetable). However, when deviations from the plan occur (for example, due to delays, cancelations, or reduction of track capacity), the dispatchers have to find alternative routes and schedules for the trains. Commonly this is done under objectives such as minimizing the number of delays, the overall delays, the number of missed connections, the deviations from the original plan/timetable, the aggregated delays, and/or the number of cancelations, or maximizing the number of trains allowed through the system. Consequently, we deal with an *optimization* problem of one sort or other. In the literature, these problem are also addressed as *Train Rescheduling* (Lüthi, Medeossi, and Nash 2009, Adenso-Diaz, Oliva

[1] See page 71.
[2] These estimates are strongly affected by the location and the extent of the construction or upgrade project.

Gonzalez, and Gonzalez-Torre 1999), *Real-Time Train Traffic Control* (Mannino 2011, Mazzarello and Ottaviani 2007)*,* *Train Rerouting* (if also routing decisions must be taken; Corman, D'Ariano, Pacciarelli, and Pranzo 2010a), *Delay Management* (Schöbel 2007), or *Traffic Disturbance Management* (Törnquist 2006) .

Train dispatchers typically have to take into account multiple (and sometimes conflicting) issues, such as

- safety – which is ensured by giving trains sufficient separation,
- avoiding deadlocks,
- neutrality to which company is given priority,
- lines/trains that may have different priorities, or
- the impact of decisions in the near future on operations later that day.

All these issues have to be dealt with by communicating with the drivers (and sometimes the guards) on board of the trains, while, at the same time, receiving information from multiple data sources.

The rescheduling decisions are of crucial relevance as they determine the response of the entire network to unpredictable deviations from the expected behavior or timetable. Actually, a decision taken in a station can affect the behavior of the traffic in another station hundreds of kilometers away. Rescheduling decisions which must be taken quickly can affect the behavior of the whole system by reverberating through the line. Nevertheless, these decisions are still mostly taken by human operators during a very limited time span. To manage this formidable task, a common solution is to divide the dispatching of trains for a larger geographical area among multiple dispatchers, each controlling a smaller region. Such a decomposition helps the dispatchers to find good solutions within their particular region. Unfortunately, this does not imply that the overall solution will be good. From the operations research and the operations management literature is well known that decomposing a larger problem into smaller sub-problems can lead to sub-optimal solutions for the larger problem. Having methods that enable dispatchers to make good and efficient plans for more trains in a greater geographical region for a longer planning horizon will make it possible to utilize recourses in a better way than splitting the problem into multiple smaller problems.

Different approaches and tools support the dispatchers; they are described in Section 2.1 of this report. It is, however, clear that the problem faced by the dispatchers becomes more and more challenging with increasing network congestion and occurrence of delays. It is also well known that delays occur more frequently in congested networks. To make things worse, it has been proven for idealized case examples that there is a correlation between capacity consumption and risk of consecutive delays (Landex 2008). This suggests that being able to efficiently deal with delays in the near future will also reduce delays later that day.

Obviously, the transport of passengers and freight is directly affected by the quality of the dispatchers' decisions. Three examples of this can be seen in the impact the dispatching of trains has on their waiting time, the variability in their travel time, and the probability for arriving late. These factors are important for travellers when choosing modes of transportation, and are, hence, critical for the competitiveness of trains. According to Samstad et al. (2010), the variability for arriving late plays an important role for passengers on shorter travels (up to 100 km), and will, therefore, significantly increase their perceived costs of travel, making alternative modes of transportation appear more competitive.

In the following, we review the current situation in a number of countries and give an overview of scientific approaches to the rescheduling problem. We also discuss briefly the gap between these scientific approaches and the practical applications. Finally, we highlight some issues connected to socio-economic evaluations and incentive systems before concluding the report.

# 2 State of the art in prioritization methods

## 2.1 Prioritization methods applied in a real life setting

On the tactical level, well-planned timetables help to prevent the occurrence of disruptions and delays, and also their geographical and spatial propagation. This is especially important for highly utilized segments of the railway network, and often additional measures have been devised to avoid conflicts and mitigate the consequences of disruptions. Many countries have a rule-based system in place for such situations. Typically, these prioritization rules apply only on infrastructure parts which have been declared over-utilized, i.e. where there is more (planned) demand for traffic than can be accommodated within the line's capacity limits in a suitable manner. Prioritization rules on the tactical level provide guidelines for assigning capacity to routes during the timetabling process as opposed to operational prioritization which supports rescheduling in the case of delays and disruptions.

In **Norway**, such prioritization rules are stated in "Fordelingsforskriften" (Samferdselsdepartementet 2003, § 7-10) and the Capacity report (Jernbaneverket 2010, ch.9), the general idea being that – if applied – the prioritization shall preserve the significance of the prioritized traffic compared to the displaced traffic. In this case, public services get highest priority, followed by framework agreements, traffic for special purposes, international freight transport, other freight transport and, finally, other passenger transport. However, the infrastructure manager can deviate from this sequence if this would lead to a better total utilization of the total infrastructure capacity. Jernbaneverket (2010) describes the following hierarchy for passenger trains (with decreasing priority): International / national transport, regional, local ("nærtransport"), suburban ("forstadstransport"), and city transport.

Similar pre-determined prioritization criteria apply, for example, in **Sweden** (Trafikverket 2012a). Sweden's Banverket also works with several operational levels ("driftsnivå", Trafikverket 2012b) indicating the state of the single areas or lines and consequences for train operations such as the need for prioritization rules.

However, as mentioned above, these rules and measures apply at the *tactical* rather than the operational level. They are not used to alleviate delays but to limit the effects of operational delays / disruptions and to provide guidelines for timetabling until the bottleneck situation (over-utilization) has been resolved.

More operations-oriented measures are applied in the **Japanese** railway system. There, high punctuality is crucial due to the country's strong dependency on rail and the high network utilization. NEA (2003) discusses reasons for the good punctuality records experienced on the Japanese railway system: For example, high standards for maintenance, including preventive maintenance of a simple and robust infrastructure, as well as simple and well-proven systems with redundancy lead to few failures in services. Management focuses on eliminating causes for disruptions, thoroughly examining occurred delays as lessons for future improvement. But also service-oriented staff with a strong work ethos and disciplined passengers contributes to the good record. Although timetables are optimized with respect to capacity utilization, a propagation of disruptions to other lines through the tight schedules is prevented by confining staff and rolling stock to a given line.

Little literature and publicly available guidelines seem available on how disruptions are actually dealt with and resolved when they first occur. Also the report on the Japanese system does not give much detail.

In **Norway**, the main rule for the dispatchers is that trains on time shall be prioritized over trains which are not on time. Among the latter, Oslo-bound trains have higher priority. However, based on the overall situation and personal experience, dispatchers can deviate from these guidelines and prioritize or cancel trains.

**Sweden** uses similar guidelines stating that trains on time shall be prioritized over trains which are not on time (Olsson, Sætermo, and Røstad 2002). It is, however, up to the different train operating companies to give suggestions for prioritization between their own trains that are delayed.

In **Denmark**, the main rule for prioritization on the operational level is that trains on time shall be prioritized over trains which are not on time. During daytime, passenger trains have priority, while freight trains have priority at night (Olsson et al. 2002). There are, however, limitations on how much a passenger train can delay a freight train and vice versa. If a train of a given type (for example, a passenger train) which is supposed to have priority is being delayed and that results in a delay of more than 15 minutes for the train with lower priority (a freight train), then the train with the lower priority is allowed to pass the delayed train.

According to Olsson et al. (2002), **Finland** has not implemented operational rules which apply to the entire country. Instead, each geographical area operates according to own rules with great flexibility for each dispatcher. Within the different types of passenger trains, high speed trains are given priority over intercity trains which again have priority over night trains.

**Austria** runs an operational prioritization scheme based on different train products (Olsson et al. 2002). The Austrian dispatchers have quite a lot of freedom to make their own judgment of what would be the best way of prioritizing trains for the system seen as a whole.

In **Italy**, dispatchers try to maximize the number of trains on time, rather than using prioritization rules to differentiate between different types of trains. It can be noted, however, that different types of trains may have different thresholds for how many minutes after their scheduled departure time they have to depart to avoid being considered delayed.

To our knowledge, rescheduling problems are typically still dealt with manually by human operators (dispatchers), aided by the aforementioned guidelines, experience and local knowledge. Hence, the performance of such human decisions varies significantly between dispatchers (Lüthi et al. 2009). However, automated systems support the dispatchers in their work.

## Train dispatching systems currently in use

Infrastructure managers use various systems to monitor and control a wide range of their operations. This section presents systems used for train dispatching. Some of these systems can be connected with other systems to control and monitor other aspects of train operations. However, we are here interested solely in how they are use in train operations to identify delayed trains and enable dispatchers to resolve conflicts by rescheduling trains.

The **ABB RailManager** system offers automatic train management (ABB 2012). According to a user of the system, this implies that the system can determine a path for each train to follow through the network (for example, given in the timetable) and has built-in procedures for conflict detection. The system is used by Jernbaneverket in Trondheim.

The **ICONIS integrated control centre** by Alstom is described as *"to guarantee train adherence to schedules, route automatic setting, resource usage in a conflict free way"*. No specification is given of a coordinated decision support option for multiple trains over a larger geographical area. The system (or parts of the system) is used in 20 different countries. In Europe it has been implemented at the following locations (Alstom 2012):

| Country | Location |
|---|---|
| **France** | Paris (RATP) |
| **Greece** | Athens (Ergose, Attico Metro) |
| **Italy** | Bologna (RFI), Rome (ATAC), Milan (FNM) and Sardinia (RFI) |
| **Spain** | Madrid (MINTRA) and Malaga (Metro Malaga) |
| **Switzerland** | Lausanne (MLO, LEB) |
| **UK** | London (Eurotunnel) |

The **SIEMENS VICOS** system offers automatic train tracking, conflict detection and automatic route setting (Siemens 2012). Conflicts are resolved automatically. The system is described to be able to automatically dispatch trains in line with operational priorities, the flow of transport, timetable adherence and connection assurance. The system is currently in use at the following locations:

| Country | Line/location |
|---|---|
| **China** | Beijing – Tianjin |
| **Denmark** | Grenå Line, Århus-Odder |
| **Estonia** | Narva |
| **Finland** | Tampere, Helsinki |
| **Greece** | Athens – Corinth – Kiato, Athens – Piraeus |
| **Lithuania** | Siauliai – Klaipeda |
| **The Netherlands** | HSL Zuid |
| **Norway** | Oslo, Drammen, Lillestrøm |
| **Romania** | Ploesiti |
| **Saudi Arabia** | Dammam – Riyadh |
| **Syria** | Deir Ezzor – Abou Kama, Tartous – Lattakia |
| **UK** | Bournemouth, Portsmouth |

As evident from this overview, a variety of automated systems has been implemented, helping to detect and resolve conflicts. In principle, automated approaches allow to consider larger-scale effects and find good rescheduling procedures faster than even an experienced human dispatcher may achieve. However, none of the systems describe options for *coordinated decision support for larger parts of the network*. Moreover, rather advanced conflict resolution and rescheduling techniques based on, e.g., sophisticated mathematical approaches have been implemented only scarcely so far. A few exceptions exist. Two implementations of automatic rescheduling systems using optimization methods are currently in operation in Europe:

1. the Lötschberg Base Tunnel system, operated by Swiss BLS (Montigel 2006),
2. a few Italian single-track railway lines (Mannino 2011) – here, still an approving interaction by a dispatcher is required before the system's decision is carried out.

Both systems appear to work quite well, but official statistics are not available. Montigel (2006) reports that the Swiss system is "… contributing to an increase of capacity and stability over the overall system, as well as gaining time for most obstructed trains". A third rescheduling system was in operation in Milano metro stations (Mannino and Mascis 2009) but has since been decommissioned. This is due to the fact that Siemens won a tender for the complete renewal of the three lines, replacing the former Bombardier controlling system. Tests over some periods proved that – on average – the system was performing substantially better than the dispatchers on average and showed a prominent increase in punctuality and regularity. In addition, a wide-spread implementation of fully optimized systems is scheduled for the end of 2012 on a few lesser-utilized regions of the Italian railway network.

## 2.2 State of the art in research

Optimization techniques allow for a formal and unambiguous description of the situation, also in a larger geographical and spatial context. This facilitates the analysis of more complex situations, taking into account knock-on effects and value estimations. Analysing the mathematical properties of the resulting model, interesting and useful properties can be derived which can help to find good solutions fast and reliably. While rule-based methods typically find *feasible* suggestions (observing all given restrictions etc.) optimization methods can determine the *best possible* (optimal) among the feasible solutions. They also can be used to evaluate the quality of several available suggestions. Typically, they can also be easily adapted to cope with new constraints or different objectives.

Rescheduling (optimization) problems can be represented and solved by means of *mathematical programming*, in particular, by *Mixed Integer Linear Programming* (MILP), where the decision variables are continuous or binary, and constraints and objectives can be represented by linear constraints. Many scientific approaches to the rescheduling problem include some kind of MILP model to describe in detail the problem's decision variables and constraints. However, general MILP solution methods are then not exploited, and, instead, the authors resort to some type of ad-hoc heuristic procedure to find a feasible solution. As a consequence, the optimality or even simply the quality of the solutions cannot be guaranteed. There are a few exceptions such as Schöbel (2007), Mannino and Mascis (2009), Sato and Fukumura (2010), or Acuna-Agost, Michelon, Feillet and Gueye (2011). Actually, when MILP solution techniques are used, they seem to always outperform other approaches (Acuna-Agost et al. 2011).

The ability of optimization to increase punctuality has been shown in several experimental studies. In these studies, input data to the optimization models has either been generated by simulation or been obtained from actual operations. The performance of these models has been evaluated using simulation. For example, Lüthi et al. (2009) report up to 100% improvement in a very congested area including Lucerne station. For the Schiphol junction in the Netherlands, Mazzarello and Ottaviani (2007) present improvements of about 3% while Corman, D'Ariano, Pacciarelli, and Pranzo (2010b) achieve an up to 12% better punctuality for the Schiphol dispatching area and up to 6% for the Utrecht dispatching area. Rodriguez (2007) studies ten realistic test instances for the Pierrefitte-Gonesse junction (North Paris) and finds improvements up to 93% (ca. 75% on average). Favourable comparisons for a case study on a single-track line in India (Surat-Bharuc) are also reported in Sundaravalli, Narayan and Rangaraj (2012).

Mannino and Mascis (2009) present a real-life implementation in a metro station in Milano. Their results show, after an 8-day test period, an average improvement of almost 9% in punctuality and regularity. The application of optimization based methods also contributes to energy savings, which becomes more significant when re-speeding is allowed; for example, Mazzarello and Ottaviani (2007) report about 10% energy savings for the Schiphol area.

Finally, efficient rescheduling procedures allow less buffer time to achieve the same level of system reliability (Lüthi et al. 2009). This provides an alternative to building new or extending existing infrastructure, reducing train headways by improving signalling systems (e.g., European Train Control System – ECTS), or reducing train headways by reducing buffer times necessary to protect from knock-on delays. For example, the latter approach increases interdependencies among trains and thus the number of potential conflicts, hence, increasing pressure on dispatchers to find good solutions if rescheduling is required.

## 2.3 Bridging the gap between theory and practice

In spite of very promising results reported in the literature, there are very few systems actually in operation which rely to a certain extent on optimization techniques. Moreover, even these exceptions are more focused on automation rather than on optimization. This fact is quite surprising as the expected benefits would be large. However, there is a number of possible motivations.

1. Early attempts to automate the dispatching process have failed. Typically, such attempts were based on expert systems or simulation, and created a general distrust in the possibility of replacing human dispatching.
2. Scientists often consider simplified models, in general not fully adherent to the real problem. This does not mean that the results become less promising or non-exploitable. However, a great effort has to be done to coordinate the real-world problems with the optimization models and algorithms developed in the academic world.
3. Signaling systems and, more generally, dispatching procedures need to be adapted.
4. Real-life instances are typically very large and some more research may be required to efficiently tackle such complexity.

Nevertheless, the interest of railway operators in the use of automatic re-schedulers is growing worldwide. This is testified by the flourishing of research projects also sponsored by rail operators. In turn, this induced also a growing interest in the academia, as testified by the number of recent Ph.D. programs devoted to this subject (de Oliveira 2001, Törnquist 2006, D'Ariano 2008, Lüthi 2009, Corman 2010). All studies converge towards a clear result: the use of some kind of automatic and possibly mathematical approach to the optimization problem significantly increases the performances of the network. But another and clearer sign of interest is that automatic and optimized rescheduling systems are explicitly mentioned or required in official documents. An example is the recent call for tender for the total renewal of the Danish signaling infrastructure (Søndergaard 2011). Analogously, real-time rescheduling and optimization is required by the operator in the renewal of the Tiburtina Station, the second largest station in Rome and one of the largest in Italy.

# 3 Socio-economic issues and incentive systems

Considering congestion in transportation, three different planning levels can be discerned. On the *long-term* or *strategic* level, decision makers typically deal with decisions about where and when to add or remove capacity in terms of infrastructure (which is beyond the scope of this report). On the *medium-term* or *tactical* level, congestion charges, the creation of timetables and other issues relevant for planning the use of the infrastructure in the coming weeks and months[3] are considered. On the *short-term* or *operative* level, the focus is on the current traffic and on reducing the impact of delays and congestion for the next few hours.

Current research at SINTEF in cooperation with Jernbaneverket, Flytoget AS and Cargolink investigates how to determine charges and subsidies for infrastructure access on the *tactical* level to achieve an efficient utilization of the Norwegian railroad infrastructure. Note that such tariffs help to shift traffic from congested periods to uncongested periods or also eliminate it. On the *operational*[4] level, the use of congestion charges has become more common in road transportation over the last few years. Such charges have been implemented in cities like Trondheim, Stockholm, London, and Singapore. The idea is that they reduce the demand for infrastructure at peak hours such that capacity is no longer exceeded and queues are reduced or eliminated. As the charges are typically levied also outside rush time, another effect is the collection of money used, e.g., for infrastructure improvement.

[3] In the Norwegian system this typically means six months, as the timetable is updated bi-annually.
[4] We define this to be on the operational level (despite the tariffs usually being set some time in advance) as travel by car typically is not planned far into the future.

The level of congestion charges is usually determined several months in advance. This means that the fees must be based on predicted levels of traffic; alternatively, they are the result of a political process and potentially decoupled from socio-economic theory. An effect of such a static system is that, as soon as traffic deviates slightly from the original estimated levels, the charges may no longer be optimal and, hence, the infrastructure may not be used to its full potential or be over-utilized. To be able to use the infrastructure system to its full potential one should either create an efficient market, determining infrastructure usage charges/subsidies by considering the infrastructure system as a whole or implement some prioritization system. Still, congestion charges have proven to reduce congestion in the cities where it has been implemented. However, such systems do generally not differentiate between the socio-economic values of the different infrastructure users. Consequently, priority will be given to those who value their time the highest and have the means to pay the tariffs. The individuals' own evaluation of what to do and their ability to pay the tariffs may not be the same as society's evaluation of who should have access to scarce infrastructure, which can result in a loss for society. Developing a prioritization scheme can create a direct link between who should get access to the infrastructure according to socio-economic principles and who is actually getting access. Such a prioritization scheme would overcome the problems associated with only relying on a regime of fees.

Obviously, the sheer amount of information needed and the complexity of the problem make it near impossible to implement operational prioritization schemes efficiently controlling access to road infrastructure on a socio-economic basis. In contrast, transportation systems such as railroad systems have few dimensions of freedom which reduces the physical possibilities to remove trains from a pre-defined sequence. In addition railroad systems involve only a limited number of users (e.g., trains) with quite predictable behaviour, each usually carrying a high number of passengers or large quantities of freight. This facilitates the use of advanced prioritization schemes. Knowing the origin, destination and intermediary stops for each train, combined with information about the type and number of passengers (or type and quantity of freight) at each station/terminal, makes it possible to assign a fairly precise socio-economic value to each train. Additionally, knowing the train's position in the system at all times, trains can be prioritized to the best of all users of the system and within all capacity limitations (typically, track and station capacity). As new and updated information becomes available every few seconds, the decision maker may have to review and frequently change these decisions. This frequent update and review facilitates good or optimal decisions over a larger time span and, thus, an efficient utilization of society's scarce resources.

## Combining the dispatching of trains and quality improving mechanisms

In January 2012, so-called quality fees were introduced in **Sweden** and shall be tested for some time (Trafikverket 2012c). These fees are to be paid by the part (train operator or infrastructure manager) responsible for a disruption to the affected parts if this disruption leads to a delay of over five minutes – excluding knock-on effects, though. They shall induce quality-increasing behaviour, ultimately leading to an increased punctuality record and reduced risk of disruptions. In 2012, the charge amounts to 10 SKR per delay minute while in 2013, it will be 15 SKR. These fees may in the future be reviewed and differentiated with respect to time of day, geography, etc. (Trafikverket 2012c, page 73).

A similar system has been implemented in December 2009 in **Germany** (Deutsche Bahn 2011). There, the percentage of trains delayed at arrival – i.e. over 6 minutes for passenger trains and 31 minutes for freight trains – is determined for each operator over a whole year. If this percentage exceeds a given threshold value, charges or bonuses are calculated according to the number of delay minutes and their causes. At present, this charge amounts to 0.10 € per minute, and only the (from a customer point of view) most important trains are considered.

Currently the incentive mechanisms mentioned above are not in use in **Norway**. However, the agreement between the Ministry of transportation and NSB AS (Samferdselsdepartementet 2011) includes clauses related to punctuality and regularity which can be interpreted as an incentive scheme.

While the above quality fees may help to improve on-time performance, they do not fully reflect the socio-economic costs associated with a given delay. In their current form, there is no differentiation between an almost empty train delayed in the middle of the night causing few or no problems for other trains and a full train delayed in the middle of rush hour causing delays for several other trains. It is likely to assume that these quality fees may be more effective if they reflected the actual cost on the system due to the delays. Given the static nature of the current implementation, it seems problematic to connect the fees to some form of prioritization rules or to use them to deal with problems in an operational (short-term) setting.

However, one a way forward to better connect prioritization methods with the actual value of each train is the following. Currently, estimates of varying quality exist on the number of passengers and the quantity of freight on board each train as well as the trains' position at all times throughout the day. Knowing the current position of each train as well as the timetable makes it easy to determine if a train is ahead of time, on time or delayed. This information allows calculating the socio-economic optimal sequence and the paths for all trains throughout the day as information becomes available. The result is a list of where and when each train should drive such that the assets are used in an optimal or near optimal way for every hour of the day. Such a list helps to give a good estimate of the socio-economic value of dispatching the trains in this particular way. One can also analyse how a certain delayed train would affect the priorities and thereby the system's socio-economic value if it would have been on time simply by comparing the priority lists and, in turn, the evaluations of socio-economic utility for these two situations. The difference between the two evaluations is the delay's socio-economic impact on the system, which again may be connected to some quality fees. It is important to note that such an evaluation does not take into account the effects of delays on operations much later that day. Neither can it be used to calculate the costs of *multiple* trains being late or cancelled.

This evaluation scheme will give estimates of the costs associated with different trains being delayed at different parts of the network at different times. However, when performing such calculations over time, it is not unlikely that some trains turn out to be delayed more often than others. Even more important, some delays may have a greater negative impact than delays of other trains. As a consequence, the train operating companies may want to take measures to improve the on-time performance of just those trains, for example, to assign a more conservative driving time or to insert other types of slack into the plans. Conversely, measures may promote trains / routes with few delays and where the delays cause only minor problems.

## Comparability to other modes of transportation

As road transportation is the obvious alternative to rail it appears sensible for decision makers to compare the two modes in an as unbiased way as possible. Being able to give optimal priorities according to socio-economic principles to a portfolio of trains through a portfolio of bottlenecks can provide information on the socio-economic value of:

1. The passengers and freight on each train at all times – including the cost of potential delays.
2. The passengers and freight between each origin-destination pair at different times of the day and over the day.

Compared to road transportation, the first aspect corresponds to determining the value of what is transported on a road segment by car and the costs associated with queues and delays. The second aspect gives a measurement of the value of what is transported on rail versus the costs associated with building and operating the infrastructure, which again can be compared to values for road infrastructure. Further, it is possible to get better measurements of the shifts in transportation between road and rail as travel times, congestion and delays fluctuate over the day. Consequently, this can support decisions about capacity improvements for both modes of transportation as well as about coordinated use of existing infrastructure.

## 4 Conclusions

We investigated measures to support traffic control on potentially highly congested railway networks, including both currently implemented systems and academic approaches. Obviously, capacity utilization and punctuality are closely linked. In order to achieve high infrastructure utilization, high on-time performance is advantageous. However, when utilization increases it becomes a daunting task to dispatch and operate trains such that punctuality and regularity remain high. Our focus in this report has, therefore, been on train rescheduling in the case of delays or disruptions, that is, on the operational level.

Most systems currently in use provide dispatchers with knowledge of the position of the trains and, commonly, simple rules for how to prioritize between trains in a safe manner. However, they typically cannot support coordination over larger parts of the infrastructure network. Consequently, existing infrastructure (including trains) is not used to its full potential. Taking into account the high costs of building and maintaining infrastructure, improving decision support for dispatchers is expected to have a high return on investment.

Several studies have shown that coordinating decisions in the network, in the spatial as well as the time dimension, can have a significant effect on on-time performance and/or the capacity utilization of the infrastructure. The importance of good train prioritization methods has also been recognized as part of the national strategy for traffic and fleet management (see page 62 in NTP (2012)) which states: "I planperioden vil det bli gjennomført et større forskningsprosjekt for å utvikle optimeringsverktøy for prioritering i trafikkstyringen."

## References

ABB, *ABB RailManager. The optimal solution for railway traffic management,* 2012. http://www05.abb.com/global/scot/scot326.nsf/veritydisplay/514cf2f9fbcd99adc12575e50040a9c5/$file/abb%20railmanager_low%20resolution.pdf

R. Acuna-Agost, P. Michelon, D. Feillet, S. Gueye, *A MIP-based local search method for the railway rescheduling problem*, Networks 57(1): 69-86, 2011

B. Adenso-Diaz, M. Oliva Gonzalez, P. Gonzalez-Torre, *On-line timetable re-scheduling in regional train services,* Transportation Research Part B 33: 387-398, 1999

Alstom Transport, *ICONIS Technology for main line network control,* 2012. http://www.alstom.com/assetmanagement/DownloadAsset.aspx?ID=07351664-9b07-494e-817e-cf11a8af1c34&version=f58c118024894748b4a4007ed8720dbc2.pdf&lang=2057

F. Corman, *Real-time Railway Traffic Management: dispatching in complex, large and busy railway networks,* Ph.D. Thesis, Technical University of Delft, 2010

F. Corman, A. D'Ariano, D. Pacciarelli, M. Pranzo, *A tabu search algorithm for rerouting trains during rail operations*, Transportation Research Part B 44: 175–192, 2010a

F. Corman , A. D'Ariano, D. Pacciarelli, M. Pranzo, *Traffic management in complex and densely used networks*, Intelligent Infrastructures 42: 374-404, 2010b

A. D'Ariano, *Improving real-time train dispatching: models, algorithms and applications*, Ph.D. Thesis, Technical University of Delft, 2008

E.S. de Oliveira, *Solving single-track railway scheduling problem using constraint programming*, Ph.D. Thesis, University of Leeds, 2001

Deutsche Bahn, *Anreizsystem Trasse der DB Netz AG*, 2011. http://fahrweg.dbnetze.com/site/shared/de/dateianhaenge/infomaterial/netz/broschuere__anreizsystem__trasse.pdf

Jernbaneverket, *Kapasitetsrapporten*, 2010

Jernbaneverket, *Høyhastighetsutredningen 2010-2012, Konklusjoner og oppsummering av arbeidet i Fase 3, Del 3*, 2012 http://www.jernbaneverket.no/no/Prosjekter/Hoyhastighetsutredningen/Nyhetsarkiv/Oppsummering-Hovedkonklusjoner/

A. Landex, *Methods to estimate railway capacity and passenger delays*. Ph.D. Thesis, Danmarks Tekniske Universitet, Department of Transport, 2008

M. Lüthi, *Improving the efficiency of heavily used railway networks through integrated real-time rescheduling*, Ph.D. Thesis, University of Zurich, 2009

M. Lüthi, G. Medeossi, A. Nash, *Structure and simulation evaluation of an integrated real-time rescheduling system for railway networks*, Network and Spatial Economics 9:103-121, 2009

C. Mannino, A. Mascis, *Optimal real-time traffic control in metro stations.* Operations Research 57: 1026-1039, 2009

C. Mannino, Real-*time traffic control in railway systems.* In: A. Caprara and S. Kontogiannis (eds.) *Proceedings of Atmos'11. OASICS 20, 2011*

M. Mazzarello, E. Ottaviani, *A traffic management system for real-time traffic optimisation in railways*, Transportation Research Part B 41: 246–274, 2007

M. Montigel, *Semi-Automatic Train Traffic Control in the New Swiss Lötschberg Base Tunnel*, IRSA-Aspect 2006. www.systransis.ch/fileadmin/2006_Paper_MM.pdf

NEA, *Best of Benchmarking: BOB Railway Case. Benchmarking Passenger Transport in Railways*, NEA R20020166/57346000/PHI/IFE, Rijswijk, The Netherlands, 2003. http://ec.europa.eu/transport/rail/research/doc/bob.pdf

NTP, *Forslag til Nasjonal transportplan 2014-2023,* 2012. http://www.ntp.dep.no/2014-2023/pdf/2012_03_01_lavopploselig_utskriftsvennlig.pdf

N. Olsson, I. F. Sætermo, C. C. Røstad, *Vurdering av regler for avvikshåndtering ved uregelmessigheter i toggangen*, SINTEF Rapport STF38 FO2821, 2002

J. Rodriguez, *A constraint programming model for real-time train scheduling at junctions*, Transportation Research Part B 41: 231–245, 2007

H. Samstad, F. Ramjerdi, K. Veisten, S. Navrud, K. Magnussen, S. Flügel, M. Killi, A. Harkjerr Halse, R. Elvik, O.S. Martin, *Den norske verdsettingsstudien – Sammendragsrapport.* TØI Report 1053/2010, 2010

Samferdselsdepartementet, *Forskrift om fordeling av jernbaneinfrastrukturkapasitet og innkreving av avgifter for bruk av det nasjonale jernbanenettet (fordelingsforskriften)*, FOR-2003-02-05-135, 2003. http://www.lovdata.no/cgi-wift/ldles?doc=/sf/sf/sf-20030205-0135.html

Samferdselsdepartementet, *Avtale om utførelse av persontransport med tog som offentlig tjeneste i 2012-2017. "TRAFIKKAVTALEN" mellom Samferdselsdepartementet og NSB AS*, 2011

K. Sato, N. Fukumura, *Freight locomotive rescheduling and uncovered train detection during disruptions*, 12th WCTR, July 11-15, 2010, Lisbon, Portugal

L. Sundaravalli, N. Rangaraj, J. Venkateswaran, *Look-Ahead based Dynamic Ranking Heuristic for Railway Rescheduling,* 2012. http://www.ieor.iitb.ac.in/files/Sundar-Webs/journal3.pdf

A. Schöbel, *Integer Programming approaches for solving the delay management problem*. In: F. Geräts, L. Kroon, A. Schöbel, D. Wagner, C. Zaroliagiis (eds.) *Algorithmic Methods for Railway Optimization* (2007), Springer

Siemens. *Efficient rail transport with the Vicos OC operations control system family,* 2012. http://www.mobility.siemens.com/mobility/global/Documents/en/rail-solutions/rail-automation/operations-control-systems/vicos-oc-en.pdf

M. Søndergaard, *The signalling programme – Denmark as front runner in Europe*, Danish Rail Conference, 2011

J. Törnquist, *Railway traffic disturbance management*, Ph.D. Thesis, Blekinge Institute of Technology, 2006

Trafikverket, *Att förklara en bana överbelastad,* 2012a. http://www.trafikverket.se/Foretag/Trafikera-och-transportera/Trafikera-jarnvag/Folja-upp-trafik/Overbelastning/Att-forklara-en-bana-overbelastad/

Trafikverket, *Banverkets driftsnivåer*, 2012b. http://www.trafikverket.se/PageFiles/9435/Banverkets%20driftsnivåer.pdf

Trafikverket, *Banavgifter för ökad kund- och samhällsnytta. Delredovisning II – Remissversion*, 2012c. http://www.trafikverket.se/PageFiles/62311/banavgifter_for_okad_kund_och_samhallsnytta_delredovisning_11_remissversion.pdf